

# LARGE DEVIATIONS FOR SYMMETRISED EMPIRICAL MEASURES

JOSE TRASHORRAS

**ABSTRACT.** In this paper we prove a Large Deviation Principle for the sequence of symmetrised empirical measures  $\frac{1}{n} \sum_{i=1}^n \delta_{(X_i^n, X_{\sigma_n(i)}^n)}$  where  $\sigma_n$  is a random permutation and  $((X_i^n)_{1 \leq i \leq n})_{n \geq 1}$  is a triangular array of random variables with suitable properties. As an application we show how this result allows to improve the Large Deviation Principles for symmetrised initial-terminal conditions bridge processes recently established by Adams, Dorlas and König.

## 1. INTRODUCTION AND RESULTS

We say that a sequence of Borel probability measures  $(P^n)_{n \geq 1}$  on a topological space  $\mathcal{Y}$  obeys a Large Deviation Principle (hereafter abbreviated LDP) with rate function  $I$  if  $I$  is a non-negative, lower semi-continuous function defined on  $\mathcal{Y}$  such that

$$-\inf_{y \in A^\circ} I(y) \leq \liminf_{n \rightarrow \infty} \frac{1}{n} \log P^n(A) \leq \limsup_{n \rightarrow \infty} \frac{1}{n} \log P^n(A) \leq -\inf_{y \in \bar{A}} I(y)$$

for any measurable set  $A \subset \mathcal{Y}$ , whose interior is denoted by  $A^\circ$  and closure by  $\bar{A}$ . If the level sets  $\{y : I(y) \leq \alpha\}$  are compact for every  $\alpha < \infty$ ,  $I$  is called a good rate function. With a slight abuse of language we say that a sequence of random variables obeys an LDP when the sequence of measures induced by these random variables obeys an LDP. For a background on the theory of large deviations, see Dembo and Zeitouni [8] and references therein.

Our framework is the following: Let  $(\Sigma, d)$  be a Polish space and  $((X_i^n)_{1 \leq i \leq n})_{n \geq 1}$  be a triangular array of  $\Sigma$ -valued random variables defined on a probability space  $(E, \mathcal{E}, \mathbb{P})$ . Let  $(\sigma_n)_{n \geq 1}$  be a sequence of random variables defined on the same probability space and such that for every  $n \geq 1$  the distribution of  $\sigma_n$  is uniform over the set  $\mathfrak{S}_n$  of all permutations of  $\{1, \dots, n\}$ . We further assume that for every  $n \geq 1$  the vector  $(X_1^n, \dots, X_n^n)$  is independent from  $\sigma_n$ . Our purpose is to investigate the LD properties of the sequence of random measures

---

*Date:* Mai 29, 2007.

1991 *Mathematics Subject Classification.* 60F10, 60J65, 81S40.

*Key words and phrases.* Large deviations, random permutations, symmetrised empirical measures, symmetrised bridge processes.

Université Paris-Dauphine, Ceremade, Place du Maréchal de Lattre de Tassigny 75775 Paris Cedex 16 France. Email: xose@ceremade.dauphine.fr.

$$\mathcal{L}^n = \frac{1}{n} \sum_{i=1}^n \delta_{(X_i^n, X_{\sigma_n(i)}^n)} \quad (1.1)$$

which we call symmetrised empirical measures.

Our interest for these measures comes from the recent work of Adams, Bru, Dorlas and König [1, 2, 3, 4] on the modeling and rigorous analysis of the Bose-Einstein condensation at positive temperature. Indeed, symmetrised measures are key ingredients in the modeling of Boson systems i.e. quantum particle systems whose wave function are invariant under permutation of the single-particle variables. For such models the thermodynamic limit of the free energy is obtained by computing the limit of the properly rescaled trace of the restriction of the Boltzmann factor  $e^{-\beta \mathcal{H}_n}$  to the subspace of symmetric wave functions. In [3] and [4] this is achieved for simple models by combining LDPs for bridge processes with symmetrised initial-terminal conditions with Feynman-Kac formula and Varadhan's lemma. Here we prove that these LD results are simple consequences of the LDP for  $(\mathcal{L}^n)_{n \geq 1}$ . In this way we give a more direct and less technical proof of the LDPs in [3, 4].

Symmetrisation of sequences or vectors of random variables was already present in the literature prior to [3, 4]. In [9] the LD behavior of the empirical measures  $\frac{1}{n} \sum_{i=1}^n \delta_{Y_i}$  of an exchangeable sequence  $\mathbf{Y} = (Y_1, \dots, Y_n, \dots)$  of random variables is investigated. The analysis relies on de Finetti's theorem which identifies the distribution of  $\mathbf{Y}$  with the mixture of distributions of sequences of independent and identically distributed random variables. In [13] an LDP for the sequence of empirical measure processes  $t \in [0, 1] \mapsto \frac{1}{n} \sum_{i=1}^{[nt]} \delta_{X_i^n}$  is obtained where the  $(X_i^n)_{1 \leq i \leq n}$  are symmetrised vectors in the sense that

$$(X_1^n, \dots, X_n^n) = \frac{1}{n!} \sum_{\sigma \in \mathfrak{S}_n} (Y_{\sigma(1)}^n, \dots, Y_{\sigma(n)}^n)$$

for a given triangular array  $((Y_i^n)_{1 \leq i \leq n})_{n \geq 1}$  of random variables such that  $\frac{1}{n} \sum_{i=1}^n \delta_{Y_i^n}$  satisfies an LDP.

Now let us describe our results more precisely. First we consider the following particular case: We are given a triangular array  $((x_i^n)_{1 \leq i \leq n})_{n \geq 1}$  of fixed elements of  $\Sigma$ , possibly with repetition, and we assume that

$$\mu^n = \frac{1}{n} \sum_{i=1}^n \delta_{x_i^n} \xrightarrow{w} \mu \in M^1(\Sigma)$$

where  $\xrightarrow{w}$  stands for weak convergence in the set  $M^1(\Sigma)$  of Borel probability measures on  $\Sigma$ . Let  $(V^n)_{n \geq 1}$  be the sequence of random measures defined on  $(E, \mathcal{E}, \mathbb{P})$  by

$$V^n = \frac{1}{n} \sum_{i=1}^n \delta_{(x_i^n, x_{\sigma_n(i)}^n)} \in M^1(\Sigma^2).$$

In order to describe the LD behavior of  $(V^n)_{n \geq 1}$  we need to introduce some notations. For every  $\nu \in M^1(\Sigma^2)$  we denote by  $\nu_1$  (resp.  $\nu_2$ ) its first (resp. second) marginal. For any two probabilities  $\rho, \nu$  on a measurable space  $(A, \mathcal{A})$  we denote by

$$H(\nu|\rho) = \begin{cases} \int_A d\nu \log \frac{d\nu}{d\rho} & \text{if } \nu \ll \rho \\ +\infty & \text{otherwise} \end{cases}$$

the relative entropy of  $\nu$  with respect to  $\rho$ . Our first result is the following

**Theorem 1.** *The sequence  $(V^n)_{n \geq 1}$  satisfies an LDP on  $M^1(\Sigma^2)$  endowed with the weak convergence topology with good rate function*

$$\mathcal{I}(\nu) = \begin{cases} H(\nu|\mu \otimes \mu) & \text{if } \nu_1 = \nu_2 = \mu \\ +\infty & \text{otherwise.} \end{cases}$$

an LDP for  $(\mathcal{L}^n)_{n \geq 1}$  follows easily

**Theorem 2.** *Suppose that  $(\mathcal{L}_1^n = \frac{1}{n} \sum_{i=1}^n \delta_{X_i^n})_{n \geq 1}$  satisfies an LDP on  $M^1(\Sigma)$  endowed with the weak convergence topology with good rate function  $\mathcal{S}$ . Then  $(\mathcal{L}^n)_{n \geq 1}$  follows an LDP on  $M^1(\Sigma^2)$  endowed with the weak convergence topology with good rate function*

$$\mathcal{J}(\nu) = \begin{cases} \mathcal{S}(\nu_1) + H(\nu|\nu_1 \otimes \nu_1) & \text{if } \nu_1 = \nu_2 \\ +\infty & \text{otherwise.} \end{cases}$$

Indeed, we can consider  $\mathcal{L}^n$  as the result of a two steps mechanism: First we pick the  $(X_i^n)_{1 \leq i \leq n}$  and, once these values are given, we pick the  $\sigma_n$  and compute the resulting  $\mathcal{L}^n$ . While this may look artificial (since  $\sigma_n$  is independent from  $(X_i^n)_{1 \leq i \leq n}$ ) it allows to identify the distribution of  $\mathcal{L}^n$  as a mixture of Large Deviations Systems (hereafter abbreviated LDS), see [7]. Then Theorem 2 follows directly from Theorem 1 thanks to Theorem 2.3 in [11]. Furthermore, by keeping in mind the mixture of LDS representation, it is an easy task to extend Theorem 2 to obtain LD results for a broad class of symmetrised bridge processes including those considered in [3, 4]. For the sake of brevity we only give a statement on the LD behavior of the empirical path measure for continuous path symmetrised bridge processes. It should be clear to the reader how other LD results in [3, 4] can be extended by following our approach.

More precisely, let  $\mathbb{R}^d$  be a Polish space and  $\mathcal{C} = \mathcal{C}([0, \beta], \mathbb{R}^d)$  be the Polish space of  $\mathbb{R}^d$ -valued continuous functions defined on the interval  $[0, \beta]$  ( $\beta > 0$  fixed) endowed with the uniform convergence metric. Let  $\xi = (\xi_s)_{s \in [0, \beta]}$  be a  $\mathcal{C}$ -valued random variable defined on  $(E, \mathcal{E}, \mathbb{P})$ . We assume that for every  $(x, y) \in (\mathbb{R}^d)^2$  the distribution of  $\xi$  conditioned on  $\{\xi_0 = x, \xi_\beta = y\}$  is well-defined and denote it by  $\mathbb{P}_{x,y}^\xi$ . We further assume that for every element  $\phi$  of the set  $\mathcal{C}_b(\mathcal{C})$  of real valued bounded continuous function defined on  $\mathcal{C}$ , the map

$$(x, y) \in (\mathbb{R}^d)^2 \mapsto \mathbb{E}_{x,y}^\xi e^{\phi(\xi)}$$

is bounded and continuous. Let  $((X_i^n)_{1 \leq i \leq n})_{n \geq 1}$  be a triangular array of  $\mathbb{R}^d$ -valued random variables defined on  $(E, \mathcal{E}, \mathbb{P})$  such that  $\frac{1}{n} \sum_{i=1}^n \delta_{X_i^n}$  satisfies an LDP on  $M^1(\mathbb{R}^d)$  endowed with the weak convergence topology, with good rate function  $\mathcal{S}$ . For every  $n \geq 1$  we consider the  $\mathcal{C}^n$ -valued vector  $(\xi^1, \dots, \xi^n)$  which distribution is given by

$$\mathbb{P}_n^{sym} = \frac{1}{n!} \sum_{\sigma \in \mathfrak{S}_n} \int_{(\mathbb{R}^d)^n} \mathbb{P}(X_1^n \in dx_1, \dots, X_n^n \in dx_n) \bigotimes_{i=1}^n \mathbb{P}_{x_i, x_{\sigma(i)}}^\xi$$

and the corresponding empirical path measure

$$L_n = \frac{1}{n} \sum_{i=1}^n \delta_{\xi^i}. \quad (1.2)$$

For every  $\mu \in M^1(\mathcal{C})$  let us denote by  $\mu_0$  (resp.  $\mu_\beta$ ) the marginal distribution at time 0 (resp. at time  $\beta$ ). By  $\mu_{0,\beta}$  we denote the joint distribution at initial and final times 0 and  $\beta$ . Finally, we introduce the rate function  $L$  defined on  $M^1(\mathcal{C})$  by

$$L(\mu) = \sup_{\phi \in \mathcal{C}_b(\mathcal{C})} \left\{ \langle \phi, \mu \rangle - \int_{(\mathbb{R}^d)^2} \mu_{0,\beta}(dx, dy) \log \mathbb{E}_{x,y}^\xi e^{\phi(\xi)} \right\}.$$

The following result is a consequence of Theorem 2 in the same way Theorem 2 is a consequence of Theorem 1.

**Theorem 3.** *The sequence  $(L_n)_{n \geq 1}$  obeys an LDP on  $M^1(\mathcal{C})$  endowed with the weak convergence topology with good rate function*

$$\mathcal{T}(\mu) = \begin{cases} \mathcal{S}(\mu_0) + H(\mu_{0,\beta} | \mu_0 \otimes \mu_\beta) + L(\mu) & \text{if } \mu_0 = \mu_\beta \\ +\infty & \text{otherwise.} \end{cases}$$

### Remarks

- 1) Theorem 1.1 in [4] is a particular case of our Theorem 3 where  $\xi$  is the Brownian motion and the  $((X_i^n)_{1 \leq i \leq n})_{n \geq 1}$  are independent and have the same distribution  $\mathbf{m}$ . There it is further assumed that  $\mathbf{m}$  has compact support. Actually our result also holds for general Polish spaces instead of  $\mathbb{R}^d$ .
- 2) Applications of Theorem 3 to mean-field interacting Bosons systems with a smooth coupling functional are possible. For examples of such applications see [3].
- 3) Finally we would like to underscore the fact that there is a 3 step mechanism involved in the definition of the symmetrised bridge processes considered here. The conciseness and simplicity of our proof follows from this remark. This fact is also well illustrated by the particular form of the rate function  $\mathcal{T}$ .

There are three Sections in the remainder of the paper. Each one is devoted to the proof of one of our results.

## 2. PROOF OF THEOREM 1

Throughout this section we are given a triangular array  $((x_i^n)_{1 \leq i \leq n})_{n \geq 1}$  of elements of  $\Sigma$ , possibly with repetition, and we assume that

$$\mu^n = \frac{1}{n} \sum_{i=1}^n \delta_{x_i^n} \xrightarrow{w} \mu.$$

For every  $n \geq 1$  we note

$$\mathcal{V}_n = \left\{ \nu \in M^1(\Sigma^2) : \exists \sigma \in \mathfrak{S}_n, \nu = \frac{1}{n} \sum_{i=1}^n \delta_{(x_i^n, x_{\sigma(i)}^n)} \right\}.$$

We introduce two triangular arrays of  $\Sigma$ -valued random variables  $((L_i^n)_{1 \leq i \leq n})_{n \geq 1}$  and  $((R_i^n)_{1 \leq i \leq n})_{n \geq 1}$  defined on a probability space  $(\Omega, \mathcal{A}, \mathbb{P})$  such that for every  $n \geq 1$  the  $2n$  random variables  $L_1^n, \dots, L_n^n, R_1^n, \dots, R_n^n$  are mutually independent and each of them is distributed according to  $\mu^n$ . The sequence of random empirical measures

$$W^n = \frac{1}{n} \sum_{i=1}^n \delta_{(L_i^n, R_i^n)} \in M^1(\Sigma^2)$$

defined on  $(\Omega, \mathcal{A}, \mathbb{P})$  has the following LD behavior

**Lemma 1.** *The sequence  $(W^n)_{n \geq 1}$  satisfies an LDP on  $M^1(\Sigma^2)$  endowed with the weak convergence topology with good rate function  $H(\nu | \mu \otimes \mu)$ .*

**Proof** Since  $\mu^n \xrightarrow{w} \mu$  we have  $\mu^n \otimes \mu^n \xrightarrow{w} \mu \otimes \mu$  (see [6], Chapter 1, Theorem 3.2). The announced result then follows from Theorem 3 in [5].  $\square$

Our strategy in proving Theorem 1 consists in comparing  $V^n$  to random measures coupled to  $W^n$ . Comparison is possible because the  $\nu \in M^1(\Sigma^2)$  such that  $\mathcal{I}(\nu) < +\infty$  can be approached in the weak convergence topology by elements of  $\mathcal{V}_n$ . Our proof of this property requires to use several metrics on  $M^1(\Sigma^2)$  compatible with the weak convergence topology. This is the reason why in Section 2.1 we give a short account on the weak convergence topology prior to the proof of our approximation result. In Section 2.2 we construct our coupling and finally in Section 2.3 we give the proper proof of Theorem 1.

**2.1. An approximation result.** We are given a Polish space  $(\Sigma, d)$ . The distance  $d$  is a priori not a bounded metric but it is topologically equivalent to the bounded metric

$$\tilde{d}(x, y) = \frac{d(x, y)}{1 + d(x, y)}$$

which makes  $\Sigma$  a separable metric space but not necessarily a complete one. The product topology on  $\Sigma^2$  is metrizable by e.g.

$$d_{2,M}((x_1, x_2), (y_1, y_2)) = \max(d(x_1, y_1), d(x_2, y_2)) \quad (2.1)$$

or

$$d_{2,+}((x_1, x_2), (y_1, y_2)) = d(x_1, y_1) + d(x_2, y_2). \quad (2.2)$$

They both make  $\Sigma^2$  a Polish space. We can also metrize the product topology on  $\Sigma^2$  with the analogues  $\tilde{d}_{2,M}$  and  $\tilde{d}_{2,+}$  of (2.1) and (2.2) built on the ground of  $\tilde{d}$ .

The weak convergence topology on  $M^1(\Sigma)$  is metrizable by the so-called dual-bounded-Lipschitz metric

$$\beta_{BL,\delta}(\rho, \nu) = \sup_{\substack{f \in C_b(\Sigma) \\ \|f\|_\infty + \|f\|_{L,\delta} \leq 1}} \left\{ \left| \int_\Sigma f d\rho - \int_\Sigma f d\nu \right| \right\} \quad (2.3)$$

where  $C_b(\Sigma)$  stands for the set of bounded continuous functions defined on  $\Sigma$ ,

$$\|f\|_\infty = \sup_{x \in \Sigma} |f(x)| \quad \text{and} \quad \|f\|_{L,\delta} = \sup_{\substack{x,y \in \Sigma \\ x \neq y}} \left| \frac{f(x) - f(y)}{\delta(x,y)} \right|,$$

where  $\delta$  is either  $d$  or  $\tilde{d}$  (see [10], Chapter 11, Theorem 11.3.3). According to Kantorovitch-Rubinstein Theorem (see [10], Chapter 11, Theorem 11.8.2) the following metric on  $M^1(\Sigma)$

$$\beta_{W,\tilde{d}}(\rho, \nu) = \inf_{\substack{Q \in M^1(\Sigma^2) \\ Q_1 = \rho \quad Q_2 = \nu}} \left\{ \int_{\Sigma^2} \tilde{d}(x,y) Q(dx, dy) \right\},$$

the so-called Wasserstein metric associated to  $\tilde{d}$  is compatible with the weak convergence topology as well. However, note that the "analogue" of  $\beta_{W,\tilde{d}}$  built on the ground of  $d$  is in general not a metric for the weak convergence topology (for an illustration of this fact see [10] p.420-421). Finally we shall denote by

$$\beta_{BL,\tilde{d}_{2,M}}(\rho, \nu) = \sup_{\substack{f \in C_b(\Sigma^2) \\ \|f\|_{2,\infty} + \|f\|_{L,\tilde{d}_{2,M}} \leq 1}} \left\{ \left| \int_{\Sigma^2} f d\rho - \int_{\Sigma^2} f d\nu \right| \right\} \quad (2.4)$$

with

$$\|f\|_{2,\infty} = \sup_{(x,y) \in \Sigma^2} |f(x,y)| \quad \text{and} \quad \|f\|_{L,\tilde{d}_{2,M}} = \sup_{\substack{(x_1,y_1), (x_2,y_2) \in \Sigma^2, \\ (x_1,y_1) \neq (x_2,y_2)}} \left| \frac{f(x_1,y_1) - f(x_2,y_2)}{\tilde{d}_{2,M}((x_1,y_1), (x_2,y_2))} \right|,$$

and

$$\beta_{W,\tilde{d}_{2,+}}(\rho, \nu) = \inf_{\substack{Q \in M^1(\Sigma^2 \times \Sigma^2) \\ Q_1 = \rho \quad Q_2 = \nu}} \left\{ \int_{\Sigma^2 \times \Sigma^2} \tilde{d}_{2,+}(x,y) Q(dx, dy) \right\},$$

two metrics on  $M^1(\Sigma^2)$  compatible with the weak convergence topology. The following lemma is a key result in the proof of Theorem 1.

**Lemma 2.** *Let  $\nu \in M^1(\Sigma^2)$  be such that  $\nu_1 = \nu_2 = \mu$ . For every  $n \geq 1$  there exists a  $\nu_n \in \mathcal{V}_n$  such that  $\nu_n \xrightarrow{w} \nu$ .*

**Proof** Let  $\nu \in M^1(\Sigma^2)$  be such that  $\nu_1 = \nu_2 = \mu$ . According to Varadarajan's Lemma (see [10], Chapter 11, Theorem 11.4.1) there exists a family  $((u_i^n, v_i^n)_{1 \leq i \leq n})_{n \geq 1}$  of elements of  $\Sigma^2$  such that

$$\gamma^n = \frac{1}{n} \sum_{i=1}^n \delta_{(u_i^n, v_i^n)} \xrightarrow{w} \nu.$$

For every  $n \geq 1$  we take  $\sigma_n, \tau_n \in \mathfrak{S}_n$  such that

$$\sum_{i=1}^n \tilde{d}(u_i^n, x_{\sigma_n(i)}^n) = \min_{\sigma \in \mathcal{V}_n} \left\{ \sum_{i=1}^n \tilde{d}(u_i^n, x_{\sigma(i)}^n) \right\}$$

and

$$\sum_{i=1}^n \tilde{d}(v_i^n, x_{\tau_n(i)}^n) = \min_{\tau \in \mathcal{V}_n} \left\{ \sum_{i=1}^n \tilde{d}(v_i^n, x_{\tau(i)}^n) \right\}.$$

We shall prove that

$$\nu^n = \frac{1}{n} \sum_{i=1}^n \delta_{(x_{\sigma_n(i)}^n, x_{\tau_n(i)}^n)} \xrightarrow{w} \nu.$$

To this end it is sufficient to prove that  $\beta_{BL, \tilde{d}_{2,M}}(\nu^n, \nu) \rightarrow 0$ . Let  $\varepsilon > 0$  be fixed. There exists an  $N_0$  such that for every  $n \geq N_0$

$$\beta_{BL, \tilde{d}_{2,M}}(\nu, \gamma^n) < \varepsilon/3. \quad (2.5)$$

Since  $\gamma_1^n \xrightarrow{w} \mu$  there exists an  $N_1$  such that for every  $n \geq N_1$  we have  $\beta_{W, \tilde{d}}(\gamma_1^n, \mu^n) < \varepsilon/3$ . In a first step we will show that due to this for every  $n \geq N_1$  we have

$$\frac{1}{n} \sum_{i=1}^n \tilde{d}(u_i^n, x_{\sigma_n(i)}^n) < \varepsilon/3. \quad (2.6)$$

In a second step we will show that (2.6) leads to

$$\beta_{BL, \tilde{d}_{2,M}}(\hat{\gamma}^n, \gamma^n) < \varepsilon/3 \quad (2.7)$$

for every  $n \geq N_1$  where

$$\hat{\gamma}^n = \frac{1}{n} \sum_{i=1}^n \delta_{(x_{\sigma_n(i)}^n, v_i^n)}.$$

Analogously to (2.7) one can prove that there exists an  $N_2$  such that for every  $n \geq N_2$

$$\beta_{BL, \tilde{d}_{2,M}}(\hat{\gamma}^n, \nu^n) < \varepsilon/3. \quad (2.8)$$

By combining (2.5, 2.7, 2.8) we obtain the announced result.

Step 1. Since  $\gamma_1^n$  and  $\mu^n$  have finite support every Borel probability measure  $Q$  on  $\Sigma^2$  such that  $Q_1 = \gamma_1^n$  and  $Q_2 = \mu^n$  is of the form

$$Q(\alpha) = \frac{1}{n} \sum_{i,j=1}^n \alpha_{i,j} \delta_{(u_i^n, x_j^n)} \quad (2.9)$$

where  $\alpha = (\alpha_{i,j})_{1 \leq i,j \leq n}$  is an  $n \times n$  bi-stochastic matrix. Conversely every  $n \times n$  bi-stochastic matrix  $\alpha$  defines through (2.9) a Borel probability measure  $Q(\alpha)$  on  $\Sigma^2$  such that  $Q(\alpha)_1 = \gamma_1^n$  and  $Q(\alpha)_2 = \mu^n$ . Actually every bi-stochastic matrix can be written as a convex combination of permutation matrices. These are  $n \times n$  matrices with a single 1 in every line and every column, all other entries being 0 (for a proof of this fact see e.g. [12], Chapter 11, Example 11.2). There is an obvious one-to-one correspondence between elements of  $\mathfrak{S}_n$  and  $n \times n$  permutation matrices. For every  $\delta \in \mathfrak{S}_n$  we shall denote by  $K_\delta$  the permutation matrix naturally associated to it. Therefore for any Borel probability measure  $Q$  on  $\Sigma^2$  such that  $Q_1 = \gamma_1^n$  and  $Q_2 = \mu^n$  i.e. any choice of the components  $(\lambda_\delta)_{\delta \in \mathfrak{S}_n}$  of the convex combination  $\alpha = \sum_{\delta \in \mathfrak{S}_n} \lambda_\delta K_\delta$  such that  $Q = Q(\alpha)$  we have

$$\int_{\Sigma^2} \tilde{d}(x, y) Q(dx, dy) = \sum_{\delta \in \mathfrak{S}_n} \lambda_\delta \left( \sum_{i=1}^n \tilde{d}(u_i^n, x_{\delta(i)}^n) \right) \quad (2.10)$$

$$\geq \sum_{i=1}^n \tilde{d}(u_i^n, x_{\sigma_n(i)}^n). \quad (2.11)$$

Hence for every  $n \geq N_1$

$$\begin{aligned} \frac{1}{n} \sum_{i=1}^n \tilde{d}(u_i^n, x_{\sigma_n(i)}^n) &= \beta_{W, \tilde{d}}(\gamma_1^n, \mu^n) \\ &< \varepsilon/3. \end{aligned}$$

which proves (2.6).

*Step 2.* For every  $f \in C_b(\Sigma \times \Sigma)$  such that  $\|f\|_{2, \infty} + \|f\|_{L, \tilde{d}_{2, M}} \leq 1$  and every  $n \geq N_1$  we have

$$\begin{aligned} \left| \int_{\Sigma^2} f d\hat{\gamma}^n - \int_{\Sigma^2} f d\gamma^n \right| &\leq \frac{1}{n} \sum_{i=1}^n \left| f(x_{\sigma_n(i)}^n, v_i^n) - f(u_i^n, v_i^n) \right| \\ &\leq \frac{1}{n} \sum_{i=1}^n \tilde{d}_{2, M}((x_{\sigma_n(i)}^n, v_i^n), (u_i^n, v_i^n)) \\ &\leq \frac{1}{n} \sum_{i=1}^n \tilde{d}(u_i^n, x_{\sigma_n(i)}^n) \\ &< \varepsilon/3 \end{aligned}$$

hence  $\beta_{BL, \tilde{d}_{2, M}}(\hat{\gamma}^n, \gamma^n) < \varepsilon/3$  for every  $n \geq N_1$ .  $\square$

**2.2. Coupled empirical measures.** To every  $n \geq 1$  and every realization of  $W^n$  we associate two elements of  $M^1(\Sigma^2)$  by

$$\beta_{W, \tilde{d}_{2, +}}(W^n, \widetilde{W}^n) = \min_{\nu \in \mathcal{V}_n} \left\{ \beta_{W, \tilde{d}_{2, +}}(W^n, \nu) \right\}, \quad (2.12)$$

and

$$\beta_{W, \tilde{d}_{2, +}}(W^n, \widehat{W}^n) = \max_{\nu \in \mathcal{V}_n} \left\{ \beta_{W, \tilde{d}_{2, +}}(W^n, \nu) \right\}, \quad (2.13)$$

In case there are several elements of  $\mathcal{V}_n$  achieving the min (resp. the max)  $\widetilde{W}^n$  (resp.  $\widehat{W}^n$ ) is picked uniformly at random among these measures.

**Lemma 3.** *For every  $n \geq 1$  the random measures  $\widetilde{W}^n, \widehat{W}^n$  and  $V^n$  are identically distributed over  $M^1(\Sigma^2)$ .*

**Proof** We shall only prove that  $\widetilde{W}^n$  and  $V^n$  are identically distributed since the proof with  $\widehat{W}^n$  and  $V^n$  works the same way. Let  $n \geq 1$  be fixed. For the sake of clarity let us first assume that there is no repetition among the  $x_1^n, \dots, x_n^n$ . In this case every  $\nu \in \mathcal{V}_n$  corresponds to a single  $\tau \in \mathfrak{S}_n$  by



$$\nu = \frac{1}{n} \sum_{i=1}^n \delta_{(x_i^n, x_{\tau(i)}^n)}. \quad (2.14)$$

Next let us consider a fixed realization  $(l_i^n, r_i^n)_{1 \leq i \leq n}$  of  $(L_i^n, R_i^n)_{1 \leq i \leq n}$  and let us denote by  $w^n$  the corresponding value for  $W^n$ . Due to the property of the minimizer in the Wasserstein distance between two atomic measures we already employed in the proof of Lemma 1, for every  $\nu \in \mathcal{V}_n$  (i.e. every  $\tau \in \mathfrak{S}_n$  according to (2.14)) there exists a  $\sigma \in \mathfrak{S}_n$  such that

$$\begin{aligned} \beta_{W, \tilde{d}_{2,+}}(w^n, \nu) &= \frac{1}{n} \sum_{i=1}^n \tilde{d}_{2,+}((l_i^n, r_i^n), (x_{\sigma(i)}^n, x_{\sigma \circ \tau(i)}^n)) \\ &= \frac{1}{n} \sum_{i=1}^n \tilde{d}(l_i^n, x_{\sigma(i)}^n) + \frac{1}{n} \sum_{i=1}^n \tilde{d}(r_i^n, x_{\sigma \circ \tau(i)}^n). \end{aligned}$$

Thus, since we are looking for the maximum over  $\sigma$  and  $\tau$ , for a fixed realization  $(l_i^n, r_i^n)_{1 \leq i \leq n}$  of  $(L_i^n, R_i^n)_{1 \leq i \leq n}$  the corresponding value of  $\tilde{w}^n$  is obtained by finding  $\eta_l$  and  $\eta_r$  such that

$$\sum_{i=1}^n \tilde{d}(l_i^n, x_{\eta_l(i)}^n) = \min_{\sigma \in \mathfrak{S}_n} \left\{ \sum_{i=1}^n \tilde{d}(l_i^n, x_{\sigma(i)}^n) \right\}$$

and

$$\sum_{i=1}^n \tilde{d}(r_i^n, x_{\eta_r(i)}^n) = \min_{\sigma \in \mathfrak{S}_n} \left\{ \sum_{i=1}^n \tilde{d}(r_i^n, x_{\sigma(i)}^n) \right\}$$

and taking

$$\tilde{w}^n = \frac{1}{n} \sum_{i=1}^n \delta_{(x_{\eta_l(i)}^n, x_{\eta_r(i)}^n)}.$$

In case several  $\eta_l$  and/or  $\eta_r$  realise the minima in the displays above, those defining  $\tilde{w}^n$  are picked among them uniformly at random. Now, remark that for every  $\gamma_l, \gamma_r \in \mathfrak{S}_n$ , observing  $(l_{\gamma_l(i)}^n, r_{\gamma_r(i)}^n)_{1 \leq i \leq n}$  has the same probability as observing  $(l_i^n, r_i^n)_{1 \leq i \leq n}$  and results in  $\gamma_l \circ \eta_l$  and  $\gamma_r \circ \eta_r$  in defining  $\tilde{w}^n$  instead of  $\eta_l$  and  $\eta_r$ . Thus if we consider  $\eta_l$  and  $\eta_r$  as random variables defining  $\tilde{W}^n$ , we see that their distribution *conditioned on*  $W^n$  is uniform over  $\mathfrak{S}_n$ . Hence  $\tilde{W}^n$  and  $V^n$  are both uniformly distributed over  $M^1(\Sigma^2)$ , thus identically distributed. This proof extends easily to the case when there are repetitions among the  $x_1^n, \dots, x_n^n$ .  $\square$

**2.3. Proof of the LD bounds.** We start the proof of the LD bounds by proving the following

**Lemma 4.** *We have:*

- (1)  $\mathcal{I}$  is a good rate function.
- (2) The sequence  $(V^n)_{n \geq 1}$  is exponentially tight.

**Proof**

- (1) Let  $\alpha \geq 0$ . We have

$$\begin{aligned}
N_\alpha &= \{\nu \in M^1(\Sigma^2) : \mathcal{I}(\nu) \leq \alpha\} \\
&= \{\nu \in M^1(\Sigma^2) : H(\nu|\mu \otimes \mu) \leq \alpha\} \cap \{\nu \in M^1(\Sigma^2) : \nu_1 = \nu_2 = \mu\}.
\end{aligned}$$

Thus, for every  $\alpha \geq 0$ ,  $N_\alpha$  is the intersection of a compact and a closed subset of  $M^1(\Sigma^2)$ , therefore it is compact.

(2) For every measurable  $A \subset M^1(\Sigma^2)$  we have

$$\begin{aligned}
\limsup_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P}(V^n \in A^c) &= \limsup_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P}(W^n \in A^c | \frac{1}{n} \sum_{i=1}^n \delta_{R_i^n} = \frac{1}{n} \sum_{i=1}^n \delta_{L_i^n} = \mu^n) \\
&\leq \limsup_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P}(W^n \in A^c) \\
&\quad - 2 \liminf_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P}(\frac{1}{n} \sum_{i=1}^n \delta_{L_i^n} = \mu^n)
\end{aligned}$$

Since  $(W^n)_{n \geq 1}$  satisfies an LDP on  $M^1(\Sigma^2)$  with a good rate function it is exponentially tight (see [8], Remark a) p.8). Thus for every  $\alpha \geq 0$  we can chose a compact set  $A_\alpha \subset M^1(\Sigma^2)$  that makes the first term in the last display smaller than  $-\alpha - 2$ . On the other hand it is clear that for every  $n \geq 1$  we have

$$\mathbb{P}(\frac{1}{n} \sum_{i=1}^n \delta_{L_i^n} = \mu^n) \geq n! \frac{1}{n^n}$$

equality corresponding to the case when there are no ties among the  $x_1^n, \dots, x_n^n$ . Thus

$$-2 \liminf_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P}(\frac{1}{n} \sum_{i=1}^n \delta_{L_i^n} = \mu^n) \leq 2$$

which completes the proof.  $\square$

**2.3.1. Proof of the lower bound.** It is sufficient in order to prove the lower bound of the LDP to prove that

$$-\mathcal{I}(\nu) \leq \liminf_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P}(V^n \in B(\nu, \varepsilon))$$

holds for every  $\nu \in M^1(\Sigma^2)$  and every  $\varepsilon > 0$ , where  $B(\nu, \varepsilon)$  stands for the open ball centered at  $\nu \in M^1(\Sigma^2)$  of radius  $\varepsilon > 0$  for the  $\beta_{W, \tilde{d}_{2,+}}$  metric. So let  $\varepsilon > 0$  and  $\nu \in M^1(\Sigma^2)$  be such that  $\mathcal{I}(\nu) < +\infty$ . In particular  $\nu_1 = \nu_2 = \mu$ . According to Lemma 2 there exists a sequence  $(\nu^n)_{n \geq 1}$  of elements of  $M^1(\Sigma^2)$  such that  $\nu^n \in \mathcal{V}_n$  and  $\nu^n \xrightarrow{w} \nu$ . According to Lemma 3 we have

$$\begin{aligned}
\mathbb{P}(V^n \in B(\nu, \varepsilon)) &= \mathbb{P}(\widetilde{W}^n \in B(\nu, \varepsilon)) \\
&\geq \mathbb{P}(\beta_{W, \tilde{d}_{2,+}}(\widetilde{W}^n, W^n) < \frac{\varepsilon}{3}, \beta_{W, \tilde{d}_{2,+}}(W^n, \nu^n) < \frac{\varepsilon}{3}, \beta_{W, \tilde{d}_{2,+}}(\nu^n, \nu) < \frac{\varepsilon}{6}) \\
&\geq \mathbb{P}(\beta_{W, \tilde{d}_{2,+}}(\nu^n, W^n) < \frac{\varepsilon}{3}, \beta_{W, \tilde{d}_{2,+}}(\nu^n, \nu) < \frac{\varepsilon}{6})
\end{aligned}$$

since it follows from the definition of  $\widetilde{W}^n$  that for every  $\nu^n \in \mathcal{V}_n$  we get

$$\beta_{W, \tilde{d}_{2,+}}(\widetilde{W}^n, W^n) \leq \beta_{W, \tilde{d}_{2,+}}(\nu^n, W^n).$$

On the other hand since  $\nu^n \xrightarrow{w} \nu$  we see that for  $n$  large enough  $\left\{ \beta_{W, \tilde{d}_{2,+}}(\nu^n, \nu) < \frac{\varepsilon}{6} \right\} = \Omega$ . Thus, for those  $n$ 's

$$\begin{aligned} \mathbb{P}(\beta_{W, \tilde{d}_{2,+}}(\nu^n, W^n) < \frac{\varepsilon}{3}, \beta_{W, \tilde{d}_{2,+}}(\nu^n, \nu) < \frac{\varepsilon}{6}) &\geq \mathbb{P}(\beta_{W, \tilde{d}_{2,+}}(\nu, W^n) < \frac{\varepsilon}{6}, \beta_{W, \tilde{d}_{2,+}}(\nu^n, \nu) < \frac{\varepsilon}{6}) \\ &\geq \mathbb{P}(\beta_{W, \tilde{d}_{2,+}}(W^n, \nu) < \frac{\varepsilon}{6}) \end{aligned}$$

Finally, it follows from Lemma 1 that

$$\begin{aligned} \liminf_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P}(V^n \in B(\nu, \varepsilon)) &\geq \liminf_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P}(\beta_{W, \tilde{d}_{2,+}}(\nu, W^n) < \frac{\varepsilon}{6}) \\ &\geq -H(\nu | \mu \otimes \mu) = -\mathcal{I}(\nu). \end{aligned}$$

**2.3.2. Proof of the upper bound.** In order to prove the upper bound of the LDP, it is sufficient to prove that it holds for compact subsets of  $M^1(\Sigma^2)$ . Indeed, since  $(V^n)_{n \geq 1}$  is an exponentially tight sequence (see Lemma 4) the full upper bound will follow from Lemma 1.2.18 in [8]. Let  $A$  be a compact subset of  $M^1(\Sigma^2)$  and let us denote by

$$A_\mu = \{\nu \in A : \nu_1 = \nu_2 = \mu\}$$

which is a compact subset of  $M^1(\Sigma^2)$  as well. Since the weak convergence topology on  $M^1(\Sigma^2)$  is compatible with the  $\beta_{W, \tilde{d}_{2,+}}$  metric, it makes  $M^1(\Sigma^2)$  a regular topological space: For every  $\nu \in A$  such that  $\nu \in A_\mu^c$  there exists  $\varepsilon_\nu > 0$  such that  $B(\nu, 2\varepsilon_\nu) \cap A_\mu = \emptyset$ . In particular  $\bar{B}(\nu, \varepsilon_\nu) \cap A_\mu = \emptyset$  where  $\bar{B}(\nu, \varepsilon)$  denotes the closed ball centered on  $\nu \in M^1(\Sigma^2)$  of radius  $\varepsilon > 0$  for the  $\beta_{W, \tilde{d}_{2,+}}$  metric. On the other hand, since  $\nu \mapsto H(\nu | \mu \otimes \mu)$  is lower semi-continuous, for every  $\nu \in A_\mu$  and every  $\delta > 0$  there exists a  $\varphi(\nu, \delta) > 0$  such that

$$\inf_{\rho \in \bar{B}(\nu, \varphi(\nu, \delta))} H(\rho | \mu \otimes \mu) \geq (H(\nu | \mu \otimes \mu) - \delta) \wedge \frac{1}{\delta}.$$

For every  $\delta > 0$  we consider the coverage

$$A \subset \left( \bigcup_{\nu \in A \cap A_\mu^c} B(\nu, \varepsilon_\nu) \right) \cup \left( \bigcup_{\nu \in A_\mu} B(\nu, \frac{\varphi(\nu, \delta)}{8}) \right)$$

from which we extract a finite coverage

$$A \subset (\bigcup_{\nu \in I_1} B(\nu, \varepsilon_\nu)) \cup \left( \bigcup_{\nu \in I_2} B(\nu, \frac{\varphi(\nu, \delta)}{8}) \right)$$

where  $I_1 \subset A \cap A_\mu^c$  and  $I_2 \subset A_\mu$  are finite sets. Then, according to Lemma 1.2.15 in [8]

$$\begin{aligned} \limsup_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P}(V^n \in A) &\leq \max \left\{ \limsup_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P}(V^n \in \bigcup_{\nu \in I_1} \bar{B}(\nu, \varepsilon_\nu) \cap A), \right. \\ &\quad \left. \limsup_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P}(V^n \in \bigcup_{\nu \in I_2} B(\nu, \frac{\varphi(\nu, \delta)}{8})) \right\}. \end{aligned}$$

For every  $\nu \in I_1$  there can not be an infinite number of elements of  $\cup_{n \geq 1} \mathcal{V}_n$  in  $\bar{B}(\nu, \varepsilon_\nu) \cap A$  for otherwise we would get  $\bar{B}(\nu, \varepsilon_\nu) \cap A_\mu \neq \emptyset$ . The first term in the max is then equal to  $-\infty$ . We are left with the second term and according to Lemmas 1,2 and 3 we have

$$\begin{aligned}
\limsup_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P}(V^n \in A) &\leq \limsup_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P}(V^n \in \cup_{\nu \in I_2} B(\nu, \frac{\varphi(\nu, \delta)}{8})) \\
&\leq \max_{\nu \in I_2} \left\{ \limsup_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P}(\widehat{W}^n \in B(\nu, \frac{\varphi(\nu, \delta)}{8})) \right\} \\
&\leq \max_{\nu \in I_2} \left\{ \limsup_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P}(\beta_{W, \tilde{d}_2, +}(\nu^n, \widehat{W}^n) < \frac{\varphi(\nu, \delta)}{4}) \right\} \\
&\leq \max_{\nu \in I_2} \left\{ \limsup_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P}(\beta_{W, \tilde{d}_2, +}(\nu, W^n) < \frac{\varphi(\nu, \delta)}{2}) \right\} \\
&\leq \max_{\nu \in I_2} \left\{ - \inf_{\rho \in B(\nu, \varphi(\nu, \delta))} H(\rho | \mu \otimes \mu) \right\} \\
&\leq \max_{\nu \in I_2} \left\{ -(H(\nu | \mu \otimes \mu) - \delta) \wedge \frac{1}{\delta} \right\} \\
&\leq \max_{\nu \in I_2} \left\{ -(\mathcal{I}(\nu) - \delta) \wedge \frac{1}{\delta} \right\} \\
&\leq - \inf_{\nu \in A} \left\{ (\mathcal{I}(\nu) - \delta) \wedge \frac{1}{\delta} \right\}.
\end{aligned}$$

By letting  $\delta \rightarrow 0$  we obtain the announced upper bound, see Remark 1.2.10 in [8].

### 3. PROOF OF THEOREM 2

As mentioned in the Introduction, in order to prove Theorem 2 it is sufficient to prove that the distribution of

$$\mathcal{L}^n = \frac{1}{n} \sum_{i=1}^n \delta_{(X_i^n, X_{\sigma_n(i)}^n)}$$

on  $M^1(\Sigma^2)$  is a mixture of LDS in the sense of [11]. For the sake of clarity we recover the notations of [11] when identifying the components of the LDS:

- $\mathcal{Z} = M^1(\Sigma^2)$  is a Polish space when endowed with the weak convergence topology.
- $\mathcal{X} = M^1(\Sigma)(= \mathcal{X}_\infty)$  is a Polish space when endowed with the weak convergence topology as well.
- For every  $n \geq 1$  we note

$$\mathcal{X}_n = \left\{ \nu \in M^1(\Sigma) : \exists (x_1, \dots, x_n) \in \Sigma^n \text{ } \nu = \frac{1}{n} \sum_{i=1}^n \delta_{x_i} \right\},$$

and according to Varadarajan's lemma for every  $\nu \in \mathcal{X}$  there exists a sequence  $(\nu^n)_{n \geq 1}$  such that for every  $n \geq 1$  we have  $\nu^n \in \mathcal{X}_n$  and  $\nu^n \xrightarrow{w} \nu$ .

- The map  $\pi : \mathcal{Z} \rightarrow \mathcal{X}$  defined by  $\pi(\nu) = \nu_1$  is continuous and surjective.
- For every  $n \geq 1$  and every  $\nu = \frac{1}{n} \sum_{i=1}^n \delta_{x_i} \in \mathcal{X}_n$ , let  $P_\nu^n$  be the distribution of  $V^n = \frac{1}{n} \sum_{i=1}^n \delta_{(x_i, x_{\sigma_n(i)})}$  under  $\mathbb{P}$ . The family  $\Pi = \{P_\nu^n, \nu \in \mathcal{X}_n, n \geq 1\}$  of finite measures on the Borel  $\sigma$ -field on  $\mathcal{Z}$  is such that for every  $n \geq 1$  and every  $\nu \in \mathcal{X}_n$

we have  $P_\nu^n(\pi^{-1}\{\nu\}^c) = 0$ .

• Let  $Q^n$  be the distribution of  $\frac{1}{n} \sum_{i=1}^n \delta_{X_i^n}$ . For every  $n \geq 1$  and every measurable  $A \subset M^1(\Sigma^2)$

$$\mathbb{P}(\mathcal{L}^n \in A) = \int_{\mathcal{X}_n} P_\nu^n(A) Q^n(d\nu).$$

All the requirements of Definition 2.1 in [11] are satisfied by our model thanks to Theorem 1. It follows from Theorem 2.3 in [11] that the sequence  $(\mathcal{L}^n)_{n \geq 1}$  obeys an LDP on  $M^1(\Sigma^2)$  endowed with the weak convergence topology with good rate function

$$\mathcal{J}(\nu) = \begin{cases} \mathcal{S}(\nu_1) + H(\nu|\nu_1 \otimes \nu_1) & \text{if } \nu_1 = \nu_2 \\ +\infty & \text{otherwise.} \end{cases}$$

#### 4. PROOF OF THEOREM 3

In order to prove Theorem 3 it is sufficient to prove that the distribution of

$$L_n = \frac{1}{n} \sum_{i=1}^n \delta_{\xi^i}$$

under

$$\mathbb{P}_n^{sym} = \frac{1}{n!} \sum_{\sigma \in \mathfrak{S}_n} \int_{(\mathbb{R}^d)^n} \mathbb{P}(X_1^n \in dx_1, \dots, X_n^n \in dx_n) \bigotimes_{i=1}^n \mathbb{P}_{x_i, x_{\sigma(i)}}^\xi.$$

is a mixture of LDS. Once again, we use the notations of [11] to identify the components of the LDS:

- $\mathcal{Z} = M^1(\mathcal{C})$  is a Polish space when endowed with the weak convergence topology.
- $\mathcal{X} = M^1(\mathbb{R}^d \times \mathbb{R}^d) (= \mathcal{X}_\infty)$  is a Polish space when endowed with the weak convergence topology as well.
- For every  $n \geq 1$  we note

$$\mathcal{X}_n = \left\{ \nu \in M^1(\mathbb{R}^d \times \mathbb{R}^d) : \exists ((s_1, a_1), \dots, (s_n, a_n)) \in (\mathbb{R}^d \times \mathbb{R}^d)^n \text{ } \nu = \frac{1}{n} \sum_{i=1}^n \delta_{(s_i, a_i)} \right\},$$

and for every  $\nu \in \mathcal{X}$  there exists a sequence  $(\nu^n)_{n \geq 1}$  such that for every  $n \geq 1$  we have  $\nu^n \in \mathcal{X}_n$  and  $\nu^n \xrightarrow{w} \nu$ .

- The map  $\pi : \mathcal{Z} \rightarrow \mathcal{X}$  defined by  $\pi(\nu) = \nu_{0,\beta}$  is continuous and surjective.
- For every  $n \geq 1$  and every  $\nu = \frac{1}{n} \sum_{i=1}^n \delta_{(s_i, a_i)} \in \mathcal{X}_n$ , let  $P_\nu^n$  be the distribution of  $L^n = \frac{1}{n} \sum_{i=1}^n \delta_{\xi^i}$  under  $\bigotimes_{i=1}^n \mathbb{P}_{s_i, a_i}$ . The family  $\Pi = \{P_\nu^n, \nu \in \mathcal{X}_n, n \geq 1\}$  of finite measures on the Borel  $\sigma$ -field on  $\mathcal{Z}$  is such that for every  $n \geq 1$  and every  $\nu \in \mathcal{X}_n$  we have  $P_\nu^n(\pi^{-1}\{\nu\}^c) = 0$ .
- Let  $Q^n$  is the distribution of  $\frac{1}{n} \sum_{i=1}^n \delta_{(X_i^n, X_{\sigma_n(i)}^n)}$ . For every  $n \geq 1$  and every measurable  $A \subset M^1(\mathcal{C})$

$$\mathbb{P}_n^{sym}(\mathcal{L}^n \in A) = \int_{\mathcal{X}_n} P_\nu^n(A) Q^n(d\nu).$$

Thus, in order to prove Theorem 3, we are left to verify that for any  $\mu \in M^1((\mathbb{R}^d)^2)$  and any sequence  $\mu^n = \frac{1}{n} \sum_{i=1}^n \delta_{(s_i, a_i)} \in \mathcal{X}_n$  such that  $\mu^n \xrightarrow{w} \mu$ , the distribution

of  $L_n = \frac{1}{n} \sum_{i=1}^n \delta_{\xi_i}$  under  $\bigotimes_{i=1}^n \mathbb{P}_{s_i, a_i}^\xi$  satisfies an LDP with good rate function  $L$ . For every  $\phi \in \mathcal{C}_b(\mathcal{C})$  we have

$$\begin{aligned} \Lambda_n(\phi) &= \log \int_{\mathcal{C}^n} e^{n\langle \phi, L_n \rangle} \bigotimes_{i=1}^n \mathbb{P}_{s_i, a_i}^\xi \\ &= \sum_{i=1}^n \log \mathbb{E}_{s_i, a_i}^\xi e^{\phi(\xi)} \end{aligned}$$

Thus

$$\begin{aligned} \Lambda(\phi) &= \lim_{n \rightarrow \infty} \frac{1}{n} \Lambda_n(\phi) \\ &= \int_{(\mathbb{R}^d)^2} \mu(dx, dy) \log \mathbb{E}_{x, y}^\xi e^{\phi(\xi)}. \end{aligned}$$

The map  $\phi \in \mathcal{C}_b(\mathcal{C}) \mapsto \Lambda(\phi)$  is Gâteaux differentiable. On the other hand the distribution of  $L_n$  under  $\bigotimes_{i=1}^n \mathbb{P}_{s_i, a_i}^\xi$  is exponentially tight as can be showed by adapting the proof of Lemma 4. We conclude thanks to Corollary 4.6.14 in [8].

**Acknowledgment** I thank Omer Adelman, Halim Doss and Jérôme Renault for their interest in part of this work. I also thank the referee for his careful reading of the manuscript.

#### REFERENCES

- [1] Stefan Adams, Jean-Bernard Bru, and Wolfgang König. Large deviations for trapped interacting brownian particles and paths. *Ann. Probab.*, 34(4):1370–1422, 2006.
- [2] Stefan Adams, Jean-Bernard Bru, and Wolfgang König. Large systems of path-repellent Brownian motions in a trap at positive temperature. *Electron. J. Probab.*, 11(19):460–485 (electronic), 2006.
- [3] Stefan Adams and Tony Dorlas. Asymptotic feynman-kac formulae for large symmetrised systems of random walks. *Ann. Inst. H. Poincaré Probab. Statist.*, to appear.
- [4] Stefan Adams and Wolfgang König. Large deviations for many brownian bridges with symmetrised initial-terminal condition. *Probab. Theory Related Fields*, to appear.
- [5] John R. Baxter and Naresh C. Jain. A comparison principle for large deviations. *Proc. Amer. Math. Soc.*, 103(4):1235–1240, 1988.
- [6] Patrick Billingsley. *Convergence of probability measures*. John Wiley & Sons Inc., New York, 1968.
- [7] D. A. Dawson and J. Gärtner. Multilevel large deviations and interacting diffusions. *Probab. Theory Related Fields*, 98(4):423–487, 1994.
- [8] Amir Dembo and Ofer Zeitouni. *Large deviations techniques and applications*, volume 38 of *Applications of Mathematics (New York)*. Springer-Verlag, New York, second edition, 1998.
- [9] I. H. Dinwoodie and S. L. Zabell. Large deviations for exchangeable random vectors. *Ann. Probab.*, 20(3):1147–1166, 1992.
- [10] R. M. Dudley. *Real analysis and probability*, volume 74 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2002. Revised reprint of the 1989 original.
- [11] M. Grunwald. Sanov results for Glauber spin-glass dynamics. *Probab. Theory Related Fields*, 106(2):187–232, 1996.
- [12] C. L. Liu. *Introduction to combinatorial mathematics*. McGraw-Hill Book Co., New York, 1968.
- [13] José Trashorras. Large deviations for a triangular array of exchangeable random variables. *Ann. Inst. H. Poincaré Probab. Statist.*, 38(5):649–680, 2002.